

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gerard, Craig  
Gerard, Norma P.  
Mackay, Charles  
Ponath, Paul D.  
Post, Theodore W.  
Qin, Shixin
- (ii) TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
ANTAGONISTS THEREOF
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
(B) STREET: Two Militia Drive  
(C) CITY: Lexington  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02173
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/375,199  
(B) FILING DATE: 19-JAN-1995
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Brook, David E.  
(B) REGISTRATION NUMBER: 22,592  
(C) REFERENCE/DOCKET NUMBER: LKS94-05A PCT
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 617-861-6240  
(B) TELEFAX: 617-861-9540

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1689 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

1689 base pairs

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATCCTTTTC	CTGGCACCTC	TGATATCCTT	TTGAAATTCA	TGTTAAAGAA	TCCCTAGGCT	60
GCTATCACAT	GTGGCATCTT	TGTTGAGTAC	ATGAATAAAT	CAACTGGTGT	GTTTTACGAA	120
GGATGATTAT	GCTTCATTGT	GGGATTGTAT	TTTTCTTCTT	CTATCACAGG	GAGAAGTGAA	180
ATGACAACCT	CACTAGATAC	AGTTGAGACC	TTTGGTACCA	CATCCTACTA	TGATGACGTG	240
GGCCTGCTCT	GTGAAAAAGC	TGATACCAGA	GCACTGATGG	CCCAGTTTGT	GCCCCCGCTG	300
TACTCCCTGG	TGTTCACTGT	GGGCCTCTTG	GGCAATGTGG	TGGTGGTGAT	GATCCTCATA	360
AAATACAGGA	GGCTCCGAAT	TATGACCAAC	ATCTACCTGC	TCAACCTGGC	CATTTCCGGAC	420
CTGCTCTTCC	TCGTCACCCT	TCCATTCTGG	ATCCACTATG	TCAGGGGGCA	TAAGTGGGTT	480
TTTGGCCATG	GCATGTGTAA	GCTCCTCTCA	GGGTTTTATC	ACACAGGCTT	GTACAGCGAG	540
ATCTTTTTCA	TAATCCTGCT	GACAATCGAC	AGGTACCTGG	CCATTGTCCA	TGCTGTGTTT	600
GCCCTTCGAG	CCCGGACTGT	CACTTTTGGT	GTCATCACCA	GCATCGTCAC	CTGGGGCCTG	660
GCAGTGCTAG	CAGCTCTTCC	TGAATTTATC	TTCTATGAGA	CTGAAGAGTT	GTTTGAAGAG	720
ACTCTTTGCA	GTGCTCTTTA	CCCAGAGGAT	ACAGTATATA	GCTGGAGGCA	TTCCACACT	780
CTGAGAATGA	CCATCTTCTG	TCTCGTTCTC	CCTCTGCTCG	TTATGGCCAT	CTGCTACACA	840
GGAATCATCA	AAACGCTGCT	GAGGTGCCCC	AGTAAAAAAA	AGTACAAGGC	CATCCGGCTC	900
ATTTTTGTCA	TCATGGCGGT	GTTTTTCATT	TTCTGGACAC	CCTACAATGT	GGCTATCCTT	960
CTCTCTTCCT	ATCAATCCAT	CTTATTTGGA	AATGACTGTG	AGCGGACGAA	GCATCTGGAC	1020
CTGGTCATGC	TGGTGACAGA	GGTGATCGCC	TACTCCCACT	GCTGCATGAA	CCCGGTGATC	1080
TACGCCTTTG	TTGGAGAGAG	GTTCCGGAAG	TACCTGCGCC	ACTTCTTCCA	CAGGCACTTG	1140
CTCATGCACC	TGGGCAGATA	CATCCCATTG	CTTCCTAGTG	AGAAGCTGGA	AAGAACCAGC	1200
TCTGTCTCTC	CATCCACAGC	AGAGCCGGAA	CTCTCTATTG	TGTTTTAGGT	AGATGCAGAA	1260
AATTGCCTAA	AGAGGAAGGA	CCAAGGAGAT	NAAGCAAACA	CATTAAGCCT	TCCACACTCA	1320
CCTCTAAAAC	AGTCCTTCAA	ACCTTCCAGT	GCAACACTGA	AGCTCTTAAG	ACACTGAAAT	1380
ATACACACAG	CAGTAGCAGT	AGATGCATGT	ACCCTAAGGT	CATTACCACA	GGCCAGGGCT	1440
GGGCAGCGTA	CTCATCATCA	ACCTAAAAAG	CAGAGCTTTG	CTTCTCTCTC	TAAAAAGAGT	1500
TACCTATATT	TTAATGCACC	TGAATGTTAG	ATAGTTACTA	TATGCCGCTA	CAAAAAGGTA	1560
AAACTTTTTA	TATTTTATAC	ATTAACCTCA	GCCAGCTATT	ATATAAATAA	AACATTTTCA	1620
CACAATACAA	TAAGTAACT	ATTTTATTTT	CTAATGTGCC	TAGTTCTTTC	CCTGCTTAAT	1680
GAAAAGCTT						1689

CCCTTTTC

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Thr	Thr	Ser	Leu	Asp	Thr	Val	Glu	Thr	Phe	Gly	Thr	Thr	Ser	Tyr	1	5	10	15
Tyr	Asp	Asp	Val	Gly	Leu	Leu	Cys	Glu	Lys	Ala	Asp	Thr	Arg	Ala	Leu	20	25	30	
Met	Ala	Gln	Phe	Val	Pro	Pro	Leu	Tyr	Ser	Leu	Val	Phe	Thr	Phe	Gly	35	40	45	
Leu	Leu	Gly	Asn	Val	Val	Val	Val	Met	Ile	Leu	Ile	Lys	Tyr	Arg	Arg	50	55	60	
Leu	Arg	Ile	Met	Thr	Asn	Ile	Tyr	Leu	Leu	Asn	Leu	Ala	Ile	Ser	Asp	65	70	75	80
Leu	Leu	Phe	Leu	Val	Thr	Leu	Pro	Phe	Trp	Ile	His	Tyr	Val	Arg	Gly	85	90	95	
His	Asn	Trp	Val	Phe	Gly	His	Gly	Met	Cys	Lys	Leu	Leu	Ser	Gly	Phe	100	105	110	
Tyr	His	Thr	Gly	Leu	Tyr	Ser	Glu	Ile	Phe	Phe	Ile	Ile	Leu	Leu	Thr	115	120	125	
Ile	Asp	Arg	Tyr	Leu	Ala	Ile	Val	His	Ala	Val	Phe	Ala	Leu	Arg	Ala	130	135	140	
Arg	Thr	Val	Thr	Phe	Gly	Val	Ile	Thr	Ser	Ile	Val	Thr	Trp	Gly	Leu	145	150	155	160
Ala	Val	Leu	Ala	Ala	Leu	Pro	Glu	Phe	Ile	Phe	Tyr	Glu	Thr	Glu	Glu	165	170	175	
Leu	Phe	Glu	Glu	Thr	Leu	Cys	Ser	Ala	Leu	Tyr	Pro	Glu	Asp	Thr	Val	180	185	190	
Tyr	Ser	Trp	Arg	His	Phe	His	Thr	Leu	Arg	Met	Thr	Ile	Phe	Cys	Leu	195	200	205	
Val	Leu	Pro	Leu	Leu	Val	Met	Ala	Ile	Cys	Tyr	Thr	Gly	Ile	Ile	Lys	210	215	220	
Thr	Leu	Leu	Arg	Cys	Pro	Ser	Lys	Lys	Lys	Tyr	Lys	Ala	Ile	Arg	Leu	225	230	235	240
Ile	Phe	Val	Ile	Met	Ala	Val	Phe	Phe	Ile	Phe	Trp	Thr	Pro	Tyr	Asn	245	250	255	

CCCTGAGGAGG

Val Ala Ile Leu Leu Ser Ser Tyr Gln Ser Ile Leu Phe Gly Asn Asp  
 260 265 270  
 Cys Glu Arg Thr Lys His Leu Asp Leu Val Met Leu Val Thr Glu Val  
 275 280 285  
 Ile Ala Tyr Ser His Cys Cys Met Asn Pro Val Ile Tyr Ala Phe Val  
 290 295 300  
 Gly Glu Arg Phe Arg Lys Tyr Leu Arg His Phe Phe His Arg His Leu  
 305 310 315 320  
 Leu Met His Leu Gly Arg Tyr Ile Pro Phe Leu Pro Ser Glu Lys Leu  
 325 330 335  
 Glu Arg Thr Ser Ser Val Ser Pro Ser Thr Ala Glu Pro Glu Leu Ser  
 340 345 350  
 Ile Val Phe  
 355

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1193 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 92..1156

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTGTGCTTAT CCGGGCAAGA ACTTATCGAA ATACAATAGA AGACCCACGC GTCCGGTTTT	60
TACTTAGAAG AGATTTTCAG GGAGAAGTGA A ATG ACA ACC TCA CTA GAT ACA	112
Met Thr Thr Ser Leu Asp Thr	
1 5	
GTT GAG ACC TTT GGT ACC ACA TCC TAC TAT GAT GAC GTG GGC CTG CTC	160
Val Glu Thr Phe Gly Thr Thr Ser Tyr Tyr Asp Asp Val Gly Leu Leu	
10 15 20	
TGT GAA AAA GCT GAT ACC AGA GCA CTG ATG GCC CAG TTT GTG CCC CCG	208
Cys Glu Lys Ala Asp Thr Arg Ala Leu Met Ala Gln Phe Val Pro Pro	
25 30 35	
CTG TAC TCC CTG GTG TTC ACT GTG GGC CTC TTG GGC AAT GTG GTG GTG	256
Leu Tyr Ser Leu Val Phe Thr Val Gly Leu Leu Gly Asn Val Val Val	
40 45 50 55	
GTG ATG ATC CTC ATA AAA TAC AGG AGG CTC CGA ATT ATG ACC AAC ATC	304
Val Met Ile Leu Ile Lys Tyr Arg Arg Leu Arg Ile Met Thr Asn Ile	
60 65 70	

1193 bp cDNA

TAC	CTG	CTC	AAC	CTG	GCC	ATT	TCG	GAC	CTG	CTC	TTC	CTC	GTC	ACC	CTT	352
Tyr	Leu	Leu	Asn	Leu	Ala	Ile	Ser	Asp	Leu	Leu	Phe	Leu	Val	Thr	Leu	
			75					80					85			
CCA	TTC	TGG	ATC	CAC	TAT	GTC	AGG	GGG	CAT	AAC	TGG	GTT	TTT	GGC	CAT	400
Pro	Phe	Trp	Ile	His	Tyr	Val	Arg	Gly	His	Asn	Trp	Val	Phe	Gly	His	
		90					95					100				
GGC	ATG	TGT	AAG	CTC	CTC	TCA	GGG	TTT	TAT	CAC	ACA	GGC	TTG	TAC	AGC	448
Gly	Met	Cys	Lys	Leu	Leu	Ser	Gly	Phe	Tyr	His	Thr	Gly	Leu	Tyr	Ser	
	105					110					115					
GAG	ATC	TTT	TTC	ATA	ATC	CTG	CTG	ACA	ATC	GAC	AGG	TAC	CTG	GCC	ATT	496
Glu	Ile	Phe	Phe	Ile	Ile	Leu	Leu	Thr	Ile	Asp	Arg	Tyr	Leu	Ala	Ile	
	120				125					130					135	
GTC	CAT	GCT	GTG	TTT	GCC	CTT	CGA	GCC	CGG	ACT	GTC	ACT	TTT	GGT	GTC	544
Val	His	Ala	Val	Phe	Ala	Leu	Arg	Ala	Arg	Thr	Val	Thr	Phe	Gly	Val	
				140					145					150		
ATC	ACC	AGC	ATC	GTC	ACC	TGG	GGC	CTG	GCA	GTG	CTA	GCA	GCT	CTT	CCT	592
Ile	Thr	Ser	Ile	Val	Thr	Trp	Gly	Leu	Ala	Val	Leu	Ala	Ala	Leu	Pro	
			155					160					165			
GAA	TTT	ATC	TTC	TAT	GAG	ACT	GAA	GAG	TTG	TTT	GAA	GAG	ACT	CTT	TGC	640
Glu	Phe	Ile	Phe	Tyr	Glu	Thr	Glu	Glu	Leu	Phe	Glu	Glu	Thr	Leu	Cys	
		170					175					180				
AGT	GCT	CTT	TAC	CCA	GAG	GAT	ACA	GTA	TAT	AGC	TGG	AGG	CAT	TTC	CAC	688
Ser	Ala	Leu	Tyr	Pro	Glu	Asp	Thr	Val	Tyr	Ser	Trp	Arg	His	Phe	His	
	185					190					195					
ACT	CTG	AGA	ATG	ACC	ATC	TTC	TGT	CTC	GTT	CTC	CCT	CTG	CTC	GTT	ATG	736
Thr	Leu	Arg	Met	Thr	Ile	Phe	Cys	Leu	Val	Leu	Pro	Leu	Leu	Val	Met	
	200				205					210					215	
GCC	ATC	TGC	TAC	ACA	GGA	ATC	ATC	AAA	ACG	CTG	CTG	AGG	TGC	CCC	AGT	784
Ala	Ile	Cys	Tyr	Thr	Gly	Ile	Ile	Lys	Thr	Leu	Leu	Arg	Cys	Pro	Ser	
				220					225					230		
AAA	AAA	AAG	TAC	AAG	GCC	ATC	CGG	CTC	ATT	TTT	GTC	ATC	ATG	GCG	GTG	832
Lys	Lys	Lys	Tyr	Lys	Ala	Ile	Arg	Leu	Ile	Phe	Val	Ile	Met	Ala	Val	
			235					240					245			
TTT	TTC	ATT	TTC	TGG	ACA	CCC	TAC	AAT	GTG	GCT	ATC	CTT	CTC	TCT	TCC	880
Phe	Phe	Ile	Phe	Trp	Thr	Pro	Tyr	Asn	Val	Ala	Ile	Leu	Leu	Ser	Ser	
		250					255					260				
TAT	CAA	TCC	ATC	TTA	TTT	GGA	AAT	GAC	TGT	GAG	CGG	AGC	AAG	CAT	CTG	928
Tyr	Gln	Ser	Ile	Leu	Phe	Gly	Asn	Asp	Cys	Glu	Arg	Ser	Lys	His	Leu	
	265					270					275					
GAC	CTG	GTC	ATG	CTG	GTG	ACA	GAG	GTG	ATC	GCC	TAC	TCC	CAC	TGC	TGC	976
Asp	Leu	Val	Met	Leu	Val	Thr	Glu	Val	Ile	Ala	Tyr	Ser	His	Cys	Cys	
	280				285					290					295	
ATG	AAC	CCG	GTG	ATC	TAC	GCC	TTT	GTT	GGA	GAG	AGG	TTC	CGG	AAG	TAC	1024
Met	Asn	Pro	Val	Ile	Tyr	Ala	Phe	Val	Gly	Glu	Arg	Phe	Arg	Lys	Tyr	
				300					305						310	

SECRET 955650

CTG CGC CAC TTC TTC CAC AGG CAC TTG CTC ATG CAC CTG GGC AGA TAC	1072
Leu Arg His Phe Phe His Arg His Leu Leu Met His Leu Gly Arg Tyr	
315 320 325	
ATC CCA TTC CTT CCT AGT GAG AAG CTG GAA AGA ACC AGC TCT GTC TCT	1120
Ile Pro Phe Leu Pro Ser Glu Lys Leu Glu Arg Thr Ser Ser Val Ser	
330 335 340	
CCA TCC ACA GCA GAG CCG GAA CTC TCT ATT GTG TTT TAGGTAGATG	1166
Pro Ser Thr Ala Glu Pro Glu Leu Ser Ile Val Phe	
345 350 355	
CAGAAAATTG CCTAAAGAGG AAGGACC	1193

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 355 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Thr Ser Leu Asp Thr Val Glu Thr Phe Gly Thr Thr Ser Tyr	
1 5 10 15	
Tyr Asp Asp Val Gly Leu Leu Cys Glu Lys Ala Asp Thr Arg Ala Leu	
20 25 30	
Met Ala Gln Phe Val Pro Pro Leu Tyr Ser Leu Val Phe Thr Val Gly	
35 40 45	
Leu Leu Gly Asn Val Val Val Val Met Ile Leu Ile Lys Tyr Arg Arg	
50 55 60	
Leu Arg Ile Met Thr Asn Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp	
65 70 75 80	
Leu Leu Phe Leu Val Thr Leu Pro Phe Trp Ile His Tyr Val Arg Gly	
85 90 95	
His Asn Trp Val Phe Gly His Gly Met Cys Lys Leu Leu Ser Gly Phe	
100 105 110	
Tyr His Thr Gly Leu Tyr Ser Glu Ile Phe Phe Ile Ile Leu Leu Thr	
115 120 125	
Ile Asp Arg Tyr Leu Ala Ile Val His Ala Val Phe Ala Leu Arg Ala	
130 135 140	
Arg Thr Val Thr Phe Gly Val Ile Thr Ser Ile Val Thr Trp Gly Leu	
145 150 155 160	
Ala Val Leu Ala Ala Leu Pro Glu Phe Ile Phe Tyr Glu Thr Glu Glu	
165 170 175	
Leu Phe Glu Glu Thr Leu Cys Ser Ala Leu Tyr Pro Glu Asp Thr Val	
180 185 190	

CCCTT" 3595660

Tyr Ser Trp Arg His Phe His Thr Leu Arg Met Thr Ile Phe Cys Leu  
195 200 205

Val Leu Pro Leu Leu Val Met Ala Ile Cys Tyr Thr Gly Ile Ile Lys  
210 215 220

Thr Leu Leu Arg Cys Pro Ser Lys Lys Lys Tyr Lys Ala Ile Arg Leu  
225 230 235 240

Ile Phe Val Ile Met Ala Val Phe Phe Ile Phe Trp Thr Pro Tyr Asn  
245 250 255

Val Ala Ile Leu Leu Ser Ser Tyr Gln Ser Ile Leu Phe Gly Asn Asp  
260 265 270

Cys Glu Arg Ser Lys His Leu Asp Leu Val Met Leu Val Thr Glu Val  
275 280 285

Ile Ala Tyr Ser His Cys Cys Met Asn Pro Val Ile Tyr Ala Phe Val  
290 295 300

Gly Glu Arg Phe Arg Lys Tyr Leu Arg His Phe Phe His Arg His Leu  
305 310 315 320

Leu Met His Leu Gly Arg Tyr Ile Pro Phe Leu Pro Ser Glu Lys Leu  
325 330 335

Glu Arg Thr Ser Ser Val Ser Pro Ser Thr Ala Glu Pro Glu Leu Ser  
340 345 350

Ile Val Phe  
355

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1116 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGGGAGAAG TGAAATGACA ACCTCACTAG ATACAGTTGA GACCTTTGGT ACCACATCCT	60
ACTATGATGA CGTGGGCCCTG CTCTGTGAAA AAGCTGATAC CAGAGCACTG ATGGCCCACT	120
TTGTGCCCCC GCTGTACTCC CTGGTGTTC A CTGTGGGCCT CTTGGGCAAT GTGGTGGTGG	180
TGATGATCCT CATAAAATAC AGGAGGCTCC GAATTATGAC CAACATCTAC CTGCTCAACC	240
TGGCCATTTC GGACCTGCTC TTCCTCGTCA CCCTTCCATT CTGGATCCAC TATGTCAGGG	300
GGCATAACTG GGTTTTTTGGC CATGGCATGT GTAAGCTCCT CTCAGGGTTT TATCACACAG	360
GCTTGATACAG CGAGATCTTT TTCATAATCC TGCTGACAAT CGACAGGTAC CTGGCCATTG	420
TCCATGCTGT GTTTGCCCTT CGAGCCCGGA CTGTCACTTT TGGTGTCTATC ACCAGCATCG	480
TCACCTGGGG CCTGGCAGTG CTAGCAGCTC TTCCTGAATT TATCTTCTAT GAGACTGAAG	540

AGTTGTTTGA AGAGACTMTT TGCAGTGCTC TTTACCCAGA GGATACAGTA TATAGCTGGA 600  
 GSSATTTCCA CACTCTGAGA ATGACCATCT TCTGTCTCGT TCTCCCTCTG CTCGTTATGG 660  
 CCATCTGCTA CACAGGAATC ATCAAAACGC TGCTGAGGTG CCCAGTAAA AAAAAGTACA 720  
 AGGCCATCCG GCTCATTTTT GTCATCATGG CGGTGTTTTT CATTTTCTGG ACACCCTACA 780  
 ATGTGGCTAT CCTTCTCTCT TSCYWYMAW YCATCTTATT TGGAAATGAC TGTGAGCGGM 840  
 MGARSMWYYK GGACCTGGTC ATGCTGGTGA CAGAGGTGAT CGCCTACTCC CACTGCTGCA 900  
 TGAACCCGGT GATCTACGCC TTTGTTGGAG AGAGGTTCCG GAAGTACCTG CGCCACTTST 960  
 TCCACAGGCA CTGCTCATG CACCTGGGCA GATACATCCC ATTCCTTCCT AGTGAGAAGC 1020  
 TGGAAAGAAC CAGCTCTGTC TCTCCATCCA CAGCAGAGCC GGAACCTCTCT ATTGTGTTTT 1080  
 AGGTAGATGC AGAAAATTGC CTAAAGAGGA AGGACC 1116

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Thr	Thr	Ser	Leu	Asp	Thr	Val	Glu	Thr	Phe	Gly	Thr	Thr	Ser	Tyr	1	5	10	15
Tyr	Asp	Asp	Val	Gly	Leu	Leu	Cys	Glu	Lys	Ala	Asp	Thr	Arg	Ala	Leu	20	25	30	
Met	Ala	Gln	Phe	Val	Pro	Pro	Leu	Tyr	Ser	Leu	Val	Phe	Thr	Val	Gly	35	40	45	
Leu	Leu	Gly	Asn	Val	Val	Val	Met	Ile	Leu	Ile	Lys	Tyr	Arg	Arg		50	55	60	
Leu	Arg	Ile	Met	Thr	Asn	Ile	Tyr	Leu	Leu	Asn	Leu	Ala	Ile	Ser	Asp	65	70	75	80
Leu	Leu	Phe	Leu	Val	Thr	Leu	Pro	Phe	Trp	Ile	His	Tyr	Val	Arg	Gly	85	90	95	
His	Asn	Trp	Val	Phe	Gly	His	Gly	Met	Cys	Lys	Leu	Leu	Ser	Gly	Phe	100	105	110	
Tyr	His	Thr	Gly	Leu	Tyr	Ser	Glu	Ile	Phe	Phe	Ile	Ile	Leu	Leu	Thr	115	120	125	
Ile	Asp	Arg	Tyr	Leu	Ala	Ile	Val	His	Ala	Val	Phe	Ala	Leu	Arg	Ala	130	135	140	
Arg	Thr	Val	Thr	Phe	Gly	Val	Ile	Thr	Ser	Ile	Val	Thr	Trp	Gly	Leu	145	150	155	160



-139-

Ala Val Leu Ala Ala Leu Pro Glu Phe Ile Phe Tyr Glu Thr Glu Glu  
165 170 175  
Leu Phe Glu Glu Thr Xaa Cys Ser Ala Leu Tyr Pro Glu Asp Thr Val  
180 185 190  
Tyr Ser Trp Xaa Xaa Phe His Thr Leu Arg Met Thr Ile Phe Cys Leu  
195 200 205  
Val Leu Pro Leu Leu Val Met Ala Ile Cys Tyr Thr Gly Ile Ile Lys  
210 215 220  
Thr Leu Leu Arg Cys Pro Ser Lys Lys Lys Tyr Lys Ala Ile Arg Leu  
225 230 235 240  
Ile Phe Val Ile Met Ala Val Phe Phe Ile Phe Trp Thr Pro Tyr Asn  
245 250 255  
Val Ala Ile Leu Leu Ser Xaa Xaa Xaa Xaa Ile Leu Phe Gly Asn Asp  
260 265 270  
Cys Glu Arg Xaa Xaa Xaa Xaa Asp Leu Val Met Leu Val Thr Glu Val  
275 280 285  
Ile Ala Tyr Ser His Cys Cys Met Asn Pro Val Ile Tyr Ala Phe Val  
290 295 300  
Gly Glu Arg Phe Arg Lys Tyr Leu Arg His Xaa Phe His Arg His Leu  
305 310 315 320  
Leu Met His Leu Gly Arg Tyr Ile Pro Phe Leu Pro Ser Glu Lys Leu  
325 330 335  
Glu Arg Thr Ser Ser Val Ser Pro Ser Thr Ala Glu Pro Glu Leu Ser  
340 345 350  
Ile Val Phe  
355

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 19
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 24
- (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TACCTGCTSA ACCTGGCCNT GGCNG

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- (ix) FEATURE:  
 (A) NAME/KEY: modified\_base  
 (B) LOCATION: 9  
 (D) OTHER INFORMATION: /mod\_base= i

- (ix) FEATURE:  
 (A) NAME/KEY: modified\_base  
 (B) LOCATION: 14  
 (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACCTGGCCNT GGCNGACCTM CTCTT

25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- (ix) FEATURE:  
 (A) NAME/KEY: modified\_base  
 (B) LOCATION: 18  
 (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACCGYTACC TGGCCATNGT CCAYGCC

27

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- (ix) FEATURE:  
 (A) NAME/KEY: modified\_base  
 (B) LOCATION: 10  
 (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGCRTGGACN ATGGCCAGGT ARCGGTC

27

ACCTGGCCNT GGCNGACCTM CTCTT

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 6
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 16
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 18
- (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

NACCANRTTG TAGGGNRNCC ARMARAG

27

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 8
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 10
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 23
- (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TGTAGGGNRN CCARMARAGR AGNARGAA

28

46607 353630

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 12
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 15
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 16
- (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAAGGCGTAG ANSANNGGGT TGASGCA

27

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 4
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 7
- (D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 8
- (D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGANSANNGG GTTGASGCAG CWGTG

25

SECRET

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(A) NAME/KEY: CDS  
(B) LOCATION: 16..48

AAGCTTCCAG CAGCC ATG GAC TAC AAG GAC GAC GAT GAC AAA GAA TTC  
Met Asp Tyr Lys Asp Asp Asp Asp Lys Glu Phe  
1 5 10

48

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Asp Tyr Lys Asp Asp Asp Asp Lys Glu Phe  
1 5 10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

TTAAGAATTC ACAACCTCAC TAGATAC

27

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

CATAGTGGAT CCAGAATG

18